Northumbria Research Link

Northumbria
University
NEWCASTLE

Frankia irregularis sp. nov., an actinobacterium unable to nodulate its original host, *Casuarina equisetifolia,* but effectively nodulate members of the actinorhizal *Rhamnales*

Imen Nouioui[1], Faten Ghodhbane-Gtari[2], Manfred Rhode[3], Vartul Sangal[4], Hans-Peter Klenk[1,] Maher Gtari[5*]

1. School of Natural and Environmental Sciences, Newcastle University, Ridley Building 2, Newcastle upon Tyne, NE1 7RU, United Kingdom.
2. Laboratoire Microorganismes et Biomolécules Actives, Faculté des Sciences de Tunis, Université Tunis El Manar, 2092 Tunis, Tunisia.
3. Central Facility for Microscopy, HZI-Helmholtz Centre for Infection Research, Inhoffenstraße 7, 38124 Braunschweig, Germany.
4. Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK
5. Institut National des Sciences Appliquées et de Technologie, Université Carthage, Centre Urbain Nord, BP 676-1080 Tunis Cedex, Tunisia.

Corresponding author: Maher Gtari maher.gtari@insat.rnu.tn

Section: Actinobacteria

Keywords: *Frankia*, symbiosis, polyphasic taxonomy

Running title: Description of *Frankia irregularis* sp. nov.

The journal's contents category (New taxa-Actinobacteria)

Abbreviations: A_2pm, diaminopimelic acid; ANI, average nucleotide identity; dDDH, digital DNA–DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; GTR, general time-reversible; ML, maximum-likelihood; MP, maximum-parsimony; MRE, maximal-relative-error; MUSCLE, Multiple Sequence Comparison by Log-Expectation; PAUP, Phylogenetic Analysis Using Parsimony; RAxML, Randomized Axelerated Maximum Likelihood; TNT, Tree analysis New Technology.

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequence and the draft genome sequence reported are MH145366 and FAOZ00000000, respectively.

Abstract

A red pigmented actinobacterium designated G2^T, forming extremely branched vegetative hyphae, vesicles and mutilocular sporangia, was isolated from *Casuarina equisetifolia* nodules. The strain failed to nodulate its original host plant but effectively nodulated members of actinorhizal *Rhamnales*. The taxonomic position of G2^T was determined using a polyphasic approach. The peptidoglycan of the strain contained *meso*-diaminopimelic acid as diagnostic diamino acid, galactose, glucose, mannose, rhamnose, ribose and xylose. Polar lipid pattern consisted of phosphatidylinositol (PI), diphosphatidylglycerol (DPG), glycophospholipids (GPL1-2), phosphatidylglycerol (PG), aminophospholipid (APL) and unknown lipids (L). The predominant menaquinones are MK-9 (H_4) and MK-9 (H_6) while the major fatty acids are *iso*-$C_{16:0}$, $C_{17:1}$ ω8c and $C_{15:0}$. The size of the genome of strain G2^T is 9.5 Mb and digital DNA G+C content is 70.9%. The 16S rRNA gene showed 97.4% to 99.5 % sequence identity with the type strains of the genus *Frankia*. Digital DNA:DNA hybridisation (dDDH) values between strains G2^T and its nearest phylogenetic neighbor *Frankia elaeagni* and *Frankia discariae* type strains were below the threshold of 70 %. Based on these results, strain G2^T (=DSM 45899^T = CECT 9038^T) is proposed to represent the type strain of a novel species *Frankia irregularis* sp. nov.

Introduction

The genus name *Frankia* was first proposed by Brunchorst (1886) [1]. It belongs to the monogeneric family *Frankiaceae* (Becking 1970 emend Zhi et al. 2009) [2-3] and the order *Frankiales* (Sen et al. 2014) [4] and encompasses soil actinobacteria best known for their facultative nitrogen-fixing symbiosis with actinorhizal plants [7]. It has been shown, based on 16S rRNA [8], *gyr*B -[9] , *gln*II [8-9] genes, 16S-23S rRNA Internal Transcribed Spacers [10], MLSA (*atp*1, *fts*Z, *dna*K, *gyr*A and *sec*A) [11] and core genomes [12] phylogenies, that the genus *Frankia* is structured in four clusters in concordance with the host plant specificity proposed by Baker [13]. *Frankia* of cluster 1 are found infective on host plants of *Alnus*, *Casuarina, Allocasuarina* and *Myricaceae,* while cluster 2 represents strains that are infective on *Coriariaceae*, *Datiscaceae*, *Dryadoideae,* and *Ceanothus*. Strains of cluster 3 are the most promiscuous and are infective on *Elaeagnaceae*, *Myricaceae*, *Colletieae* and *Gynmmostoma*. The fourth *Frankia* cluster consists of the atypical strains which are unable to fix nitrogen and/or to re-infect actinorhizal host plants. Recently ten species have been recognized *Frankia alni*, *Frankia casuarinae* [14] and *Frankia canadensis* [15] of cluster 1, *Frankia coriariae* [16], *Candidatus* Frankia datiscae [17] and *Candidatus* Frankia californiensis [18] of cluster 2, *Frankia elaeagni* [14] and *Frankia discariae* (19) from cluster 3, *Frankia inefficax* [20], *Frankia asymbiotica* [21] and *Frankia saprophytica* [22] from cluster 4.

Strain G2^T of phylogenetic cluster 3, was isolated from *Casuarina equisetifolia* and appears to infect members of the *Rhamnales* order but not its original host plant. Based on a polyphasic approach, G2^T emerges as type strain of a new species *Frankia irregularis* sp. nov.

Strain G2^T was isolated from nodules collected in the INRA Research Station, Saint-François, Grande Terre, Guadeloupe [23]. The type strains of *Frankia alni, Frankia casuarinae, Frankia elaeagni* and *Frankia discariae, Frankia inefficax, Frankia asymbiotica, Frankia saprophytica, Frankia Canadensis*, *Frankia coriaria* together with the studied strain G2^T were maintained in Basic Propionate (BAP)[24] broth medium supplemented with NH4Cl at 28°C without shaking as previously described [14]. Phenotypic characterization was performed on 4 weeks old cultures. Freeze dried cells were used for chemotaxonomic analyses while a fresh wet biomass were examined for fatty acids profile and biochemical and morphological features. In this context, scanning electron microscope (FE-SEM Merlin, Zeiss, Germany) and GENIII microplates in an Omnilog device (Biolog Inc., Haywood, USA) were used as described by Nouioui *et al.* [14]. All analysed tests were carried out in duplicate.

Red pigmented colonies were developed after 3-4 weeks incubation of the type strain in BAP broth medium at 28ºC without shaking. The colonies were formed with extremely branched vegetative hyphae, vesicles and mutilocular sporangia as shown in Fig. 1a-b, features observed as well for *F. elaeagni* DSM 46783^T and *F discariae* DSM 46785^T. The ability of the type strain to fix atmospheric nitrogen and to nodulate member of the order *Rhamnales* were examined by Diem *et al.* [23]. It has been shown that strain G2^T was unable to re-infect its host plant *Casuarina equisetifolia* [23]. The type strain can be distinguished from its nearest phylogenetic neighbours, *F. elaeagni* DSM 46783^T, by its red pigmentation and several biochemical properties including its ability to metabolise bromo-succinic acid, guanidine hydrochloride, methyl pyruvate, potassium tellurite and 1% sodium lactate, and to grow in presence of minocycline and vancomycin. Moreover, strain G2^T was unable to oxidise D-glucose-6-phosphate, D-fructose-6-phosphate and *β*-hydroxy-butyric acid unlike its phylogenetic neighbour (Table 1). Thus morphological, physiological and cultural traits of strain G2^T are consistent with the genus *Frankia.*

Chemotaxonomic traits of strain G2^T have been determined based on thin layer chromatography procedures. Menaquinones and polar lipid profiles as well as diaminopimelic acids and sugars contents of whole cell hydrolysate were identified following the same protocols used by Nouioui *et al.* [14]. Fatty acid methyl esters (FAMEs) analyses for strain G2^T and the reference strains cited above were extracted and identified following the modified protocol of Miller (1982) [25] by Kuykendall *et al.* [26] and as described by Nouioui et *al*. [14]. Strain G2^T was characterized by the presence of (i) *meso*-A_2pm, galactose, glucose, mannose, rhamnose, ribose and xylose in its whole cell hydrolysates, (ii) isoprenologue profile consisted of MK-9(H_4) and MK-9(H_6) as the predominant ones (>20%) and by (iii) polar lipid pattern consisted of phosphatidylinositol (PI), diphosphatidylglycerol (DPG), glycophospholipids (GPL1-2), phosphatidylglycerol (PG), aminophospholipid (APL) and uncharacterized lipids (L). Apart from the presence of APL, the chemotaxonomic features of strain G2^T are in line with those of the type species of the genus, *Frankia alni* DSM 45986^T, and with its nearest phylogenetic neighbors; *F. elaeagni* DSM 46783^T and *F. discariae* DSM 46785^T excepting that MK-9 (H_4) was the major menaquinone and lacks rhamnose in cell wall sugars in *F. discariae* DSM 46785^T. In addition that the strain G2T contained APLThe major fatty acids (>15%) of the type strain are *iso*-$C_{16:0}$, $C_{17:1}$ ω8c and $C_{15:0}$ while the type strains of *F. elaeagni* and *F. discariae* species have $C_{16:0}$ instead of the $C_{15:0.}$

The almost complete sequences of 16S rRNA gene of strain G2^T extracted from the draft genome and obtained from PCR-product are 100% identical to each other. Pairwise 16S rRNA gene sequence similarities and phylogenetic trees were determined using the GGDC web server [27] and according to Meier-Kolthoff *et al.* [28]. Maximum-likelihood (ML) and maximum-parsimony (MP) trees were inferred on DSMZ phylogenomic pipeline [29] and using the GTR+GAMMA model. For ML and MP trees, the sequences were aligned using RAxML [30] and TNT [31], respectively. Rapid bootstrapping in conjunction with the autoMRE bootstopping criterion [32] was used for ML while 1000 bootstrapping replicates in conjunction with tree-bisection-and-reconnection branch swapping and ten random sequence addition replicates was used for MP. Multiple sequence alignment were determined using MUSCLE program [33] while X^2 test as implemented in PAUP [34] was used to check the sequence for a compositional bias.

Pairwise sequence similarities for 16S rRNA gene sequence between strain G2^T and the type strains of *Frankia* species varied from 97.4% to 99.5%. The highest values (above 99.0%) have been found with the type strains of *F. elaeagni* and *F. discariae* species which belong, with strain G2^T, to cluster 3 of the genus *Frankia* [9-10, 35-36]. In the ML phylogenetic tree, strain G2^T appeared, in highly supported clade, closely related to the type strain of *F. elaeagni* species (99.5%) forming a subclade next to the one encompasses the type strain of *F. discariae* (99.4%), *F. saprophityca* (98.1%), *F. inefficax* (97.6%) and *F. asymbiotica* (97.8%) (Fig. 2a).

The genome sequences of the *Frankia* and representative strains from other related genera were annotated using Prokka v1.11 [37] and were compared using BPGA 1.3 pipeline [38]. The missing data and poorly aligned regions from concatenated protein sequence alignment of the core genome were removed using Gblocks [39]. A ML tree was constructed from the resulting alignment of 10,491 amino acids using LG+F+G4 substitution model by IQ-Tree with 100,000 ultrafast bootstrap iterations and SH-like approximate likelihood ratio tests [40]. Another ML tree was generated using PhyloPhlAn [41] which extracts subsets of amino acid sequences from 400 universal proteins and calculate phylogeny from the concatenated alignment using RAxML [42]. This approach is particularly suitable for an accurate determination of taxonomic relationships from the genomic data [41]. The phylogenetic position of strain G2^T (Fig2b and Fig2.c) is in concordance with the ML 16SrRNA gene tree

Digital DNA:DNA hybridisation (dDDH) between strain G2^T and its nearest phylogenetic neighbour cited above was calculated using genome to genome distance calculator with formula

2 available at DSMZ server (http://ggdc.dsmz.de/distcalc2.php). Strain G2^T and its phylogenetic relatives cited above showed dDDH values below the threshold of 70% designed by Wayne *et al*. [43] for delineation a novel prokaryotic species (Table 2). Strain G2^T has a genome size of 9.5 Mb with 70.9 % of G+C content while its nearest neighbours, *F. elaeagni* DSM 46783^T and *F. discariae* DSM 46785^T have respectively 7.6 Mb and 7.9 Mb with 71.7 % and 72.4 % of G+C content.

It can be concluded from the wealth of the present pholyphasic study that strain G2^T has phenotypic and genetic features consistent with those of the genus *Frankia* and distinguishable from the other *Frankia* species. Therefore, strain G2^T forms a new lineage of the genus and merits to be recognised as a new species within the genus for which the name *Frankia irregularis* sp. nov. is proposed.

Description of *Frankia irregularis* sp. nov.

Frankia irregularis (ir.re.gu.la'ris. L. fem. adj. *irregularis* of irregular, referring to the inability of the species to infect its original host plant and to infect taxonomically disparate host plants) Nitrogen fixing Gram-positive aerobic, heterotrophic and chemoorganotrophic actinobacterium known by its red pigmentation; colonies were formed by three cell structures: substrate hyphae, multilocular sporangia and vesicles. Optimal growth was observed on BAP medium for 3-4 weeks at 28ºC and from pH 6.3 to 6.8. It is able to oxidise D-cellobiose, *α-keto*-butyric acid, methyl pyruvate, L-lactic acid, bromo-succinic acid, acetic acid, guanidine hydrochloride; growth in presence of 1% sodium lactate, potassium tellurite, lincomycin, minocycline and vancomycin. Whole cell hydrolysates are formed by meso diaminopimelic acid, galactose, glucose, mannose, rhamnose, ribose and xylose; polar lipid pattern consisted of phosphatidylinositol (PI), diphosphatidylglycerol (DPG), glycophospholipids (GPL1-2), phosphatidylglycerol (PG), aminophospholipid (APL) and unknown lipids (L) (Fig. S1) and predominant menaquinones (>20%) are MK-9 (H_4) and MK-9 (H_6). The major fatty acids (>15%) are *iso*-$C_{16:0}$, $C_{17:1}$ ω8c and $C_{15:0}$.

The type strain G2^T (=DSM 45899^T = CECT 9038^T) was isolated from *Casuarina equisetifolia* [23]. The size of the genome is 9.5 Mb and digital DNA G+C content is 70.9%.

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequence and draft genome sequence reported are MH145366 and FAOZ00000000, respectively.

Funding information this work received no specific grant from any funding agency.

Conflicts of interest Authors have no conflict of interest to declare.

Acknowledgements This work was supported by Tunisian Ministry of Higher Education and Scientific Research. We are grateful to Marlen Jando and Gabriele Pötter (both at DSMZ) for their help with the chemotaxonomic analyses.

References

1. Brunchorst. J. Über einige Wurzelanschwellungen, besonders diejenigen von Alnus und den Elaegnaceen. *Botanische Institut Tubingen* 1886;2:151-77.
2. Becking JH. *Frankiaceae* fam. nov. (Actinomycetales) with one new combination and six new species of the genus Frankia Brunchorst 1886, 174. *Int J Syst Bacteriol* 1970;20:201-20.
3. Zhi XY, Li WJ, Stackebrandt E. An update of the structure and 16S rRNA gene sequence based definition of higher ranks of the class *Actinobacteria*, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa. *Int J Syst Evol Microbiol* 2009;59:589-608.
4. Sen A, Daubin V, Abrouk D, Gifford I, Berry AM, *et al*. Phylogeny of the class *Actinobacteria* revisited in the light of complete genomes. The orders 'Frankiales' and *Micrococcales* should be split into coherent entities: proposal of *Frankiales* ord. nov., *Geodermatophilales* ord. nov., *Acidothermales* ord. nov. and *Nakamurellales* ord. nov. *Int J Syst Evol Microbiol* 2014;64(Pt 11):3821-3832.
5. Benson DR, Silvester WB. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. *Microbiol Rev* 1993;57:293-319.
6. Schwencke J, Carú M. Advances in actinorhizal symbiosis: host plant-*Frankia* interactions, biology, and applications in arid land reclamation. *Arid Land Res Manag* 2001; 15:285–327.
7. Chaia EE, Wall L, Huss-Danell K. Life in soil by actinorhizal root nodule endophyte *Frankia*. *A review Symbiosis* 2010; 51:201–226.
8. Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, *et al*. Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. *Int J Syst Bacteriol* 1996;46:1-9.
9. Nouioui I, Ghodhbane-Gtari F, Beauchemin NJ, Tisa LS, Gtari M. Phylogeny of members of the *Frankia* genus based on *gyr*B, *nif*H and *gln*II sequences. *Antonie Van Leeuwenhoek* 2011;100:579-587.
10. Ghodhbane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M. 16S-23S rRNA intergenic spacer region variability in the genus *Frankia*. *Microb Ecol* 2010;60:487-495.
11. Gtari M, Ghodhbane-Gtari F, Nouioui I, Ktari A, Hezbri K, et al. Cultivating the uncultured: growing the recalcitrant cluster-2 *Frankia* strains. *Sci Rep* 2015;5:13112.
12. Louis ST, Oshone R, Sarkar I, Ktari A, Sen A, *et al*. Genomic approaches toward understanding the actinorhizal symbiosis: an update on the status of the *Frankia* genomes. *Symbiosis* 2016; 70:5-16.
13. Baker DD. Relationships among pure cultured strains of *Frankia* based on host specificity. *Physiol Plant* 1987;70:245–248.
14. Nouioui I, Ghodhbane-Gtari F, Montero-Calasanz MD, Göker M, Meier-Kolthoff JP, *et al*. Proposal of a type strain for *Frankia alni* (Woronin 1866) Von Tubeuf 1895, emended description of *Frankia alni*, and recognition of *Frankia casuarinae* sp. nov. and *Frankia elaeagni* sp. nov. *Int J Syst Evol Microbiol* 2016;66:5201-5210.
15. Normand N, Nouioui I, Pujic P, Fournier P, Dubost A, *et al*. *Frankia canadensis* sp. nov., isolated from root nodules of Alnus incana subspecies rugosa growing in Canada. *Int J Evol Syst Microbiol* 2018;Accepted.

16. Nouioui I, Ghodhbane-Gtari F, Rohde M, Klenk HP, Gtari M. *Frankia coriariae* sp. nov., an infective and effective microsymbiont isolated from *Coriaria japonica*. *Int J Syst Evol Microbiol* 2017;67:1266–1270
17. Persson T, Benson DR, Normand P, Vanden Heuvel B, Pujic P, *et al*. Genome sequence of "Candidatus Frankia datiscae" Dg1, the uncultured microsymbiont from nitrogen-fixing root nodules of the dicot *Datisca glomerata. J Bacteriol* 2011;193:7017-7018.
18. Normand P, Nguyen TV, Battenberg K, Berry AM, Heuvel BV, *et al*. Proposal of 'Candidatus Frankia californiensis', the uncultured symbiont in nitrogen-fixing root nodules of a phylogenetically broad group of hosts endemic to western North America. *Int J Syst Evol Microbiol* 2017;67:3706-3715.
19. Nouioui I, Montero-Calasanz MDC, Ghodhbane-Gtari F, Rohde M, Tisa LS, *et al*. *Frankia discariae* sp. nov.: an infective and effective microsymbiont isolated from the root nodule of *Discaria trinervis*. *Arch Microbiol* 2017;199: 641–647.
20. Nouioui I, Ghodhbane-Gtari F, Montero-Calasanz MDC, Rohde M, Tisa LS, *et al*. *Frankia inefficax* sp. nov., an actinobacterial endophyte inducing ineffective, non nitrogen-fixing, root nodules on its actinorhizal host plants. *Antonie Van Leeuwenhoek* 2017;110:313-320.
21. Nouioui I, Gueddou A, Ghodhbane-Gtari F, Rhode M, Gtari M, *et al*. *Frankia asymbiotica* sp. nov., a non infective actinobacterium isolated from Morella californica root nodule. *Int J Syst Evol Microbiol* 2017; 67:4897-4901.
22. Nouioui I, Ghodhbane-Gtari F, Klenk HP, Gtari M. *Frankia saprophytica* sp. nov. an atypical non-infective (Nod–) and non-nitrogen fixing (Fix–) actinobacterium isolated from *Coriaria nepalensis* root nodules. *Int J Syst Evol Microbiol* 2018;68:1090–1095.
23. Diem HG, Gauthier D, Dommergues YR. Isolation of *Frankia* from nodules of *Casuarina equisetifolia*. *Can J Microbiol* 1982;28:526–530.
24. Murry MA, Fontaine MS, Torrey JG. Growth kinetics and nitrogenase induction in Frankia sp. HFP ArI3 grown in batch culture. *Plant Soil* 1984;78:61-78.
25. Miller LT. Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids. *J Clin Microbiol* 1982;16:584-586.
26. Kuykendall LD, Roy MA, O'Neill JJ, Devine TE. Fatty Acids, Antibiotic Resistance, and Deoxyribonucleic Acid Homology Groups of *Bradyrhizobium japonicum*. *Int J Syst Evol Microbiol* 1988;38:358-361.
27. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. 2013;14:60.
28. Meier-Kolthoff JP, Göker M, Spröer C, Klenk HP. When should a DDH experiment be mandatory in microbial taxonomy? *Arch Microbiol* 2013;195:413-418.
29. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, *et al*. Complete genome sequence of DSM 30083(T), the type strain (U5/41(T)) of *Escherichia coli,* and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci* 2014;9:2.
30. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30(9):1312-1313.
31. Goloboff PA, Farris JS, Nixon KC. TNT, a free program for phylogenetic analysis. *Cladistics* 2008;24:774-786.
32. Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol* 2010;17:337-354.
33. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792-1797.
34. Swofford DL. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0. Sunderland: Sinauer Associates; 2002.
35. Gtari M, Brusetti L, Skander G, Mora D, Boudabous A, *et al*. Isolation of Elaeagnus-compatible Frankia from soils collected in Tunisia. *FEMS Microbiol Lett* 2004;234:349-355.
36. Gtari M, Daffonchio D, Boudabous A. Assessment of the genetic diversity of *Frankia* microsymbionts of *Elaeagnus angustifolia* L. plants growing in a Tunisian date-palm oasis by analysis of PCR amplified *nif*D-K intergenic spacer. *Can J Microbiol* 2007;53:440-445.

37. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;15:30:2068-2069.
38. Chaudhari NM, Gupta VK, Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline. *Sci Rep* 2016; 6: 24373.
39. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000;17:540-52.
40. Minh BQ, Nguyen MA, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 2013;30:1188-1195.
41. Segata N, Börnigen D, Morgan XC, Huttenhower C. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun* 2013;4:2304.
42. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006;22:2688-2690.
34. Wayne LG, Brenner BD, Colwell RR, Grimont PAD, Kandler O, *et al*. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst bacteriol* 1987; 37:463–464.

Figure Legends

Figure 1. Scanning electron (a) and light microscopy micrograph (b) of strain G2^T grew on BAP media for 4 weeks at 28ºC. (h), hyphae; (v), vesicles and (s), sporangia.

Figure 2. Maximum-likelihood phylogenetic tree based on almost complete16S rRNA gene sequences constructed using the GTR+GAMMA model. The numbers above the branches are bootstrap support values greater than 60% for ML (left) and MP (right) (a). Maximum-likelihood phylogenomic tree based on core genome sequences (b). Maximum-likelihood phylogenomic tree based on concatenated amino acid sequences from 400 universal proteins (c).

Table 1. Phenotypic and chemotaxonomic properties that distinguish strain G2^{T}from the type strains of *F. alni* DSM 45986^T, *F. asymbiotica* DSM 100626^T, *F. canadensis* DSM 45898^T, *F. casuarinae* DSM 45818^T, *F. coriariae* DSM 100624^T, *F. elaeagni* DSM 46783^T, *F. discariae* DSM 46785^T, *F. inefficax* DSM 45817^T and *F. saprophytica* DSM 105290^T. All phenotypic data was obtained in the present study.

	G2^T	DSM 45986^T	DSM 100626^T	DSM 45898^T	DSM 45818^T	DSM 100624^T	DSM 46783^T	DSM 46785^T	DSM 45817^T	DSM 105290^T
Colony colour	red	white	white	white	white	brown	red	yellow	white	white greyish
Vesicles/N_2-fixation	+	+	+	+	+	+	+	+	-	-
Carbon source										
Dextrin	-	-	-	+	-	-	-	-	+	-
D-cellobiose	+	-	+	-	-	+	+	+	-	+
β-gentiobiose	-	-	-	-	+	-	-	-	-	+
D-glucose-6-phosphate	-	+	-	+	-	-	+	+	-	+
D-fructose-6-phosphate	-	+	-	+	-	+	+	+	+	+
α-hydroxy-butyric acid	w	-	+	-	+	+	-	+	-	-
β-hydroxy-butyric acid	-	-	+	-	-	-	+	-	-	-
α-keto-butyric acid	+	-	+	+	+	-	+	-	-	+
Aceto-acetic acid	w	-	-	+	+	+	-	-	-	+
Methyl pyruvate	+	+	+	-	+	+	-	-	-	+
L-lactic acid	+	-	-	+	-	+	+	+	-	-
L-malic acid	w	+	-	-	-	-	-	-	-	+
D-malic acid	w	-	-	-	-	-	-	-	+	+
Citric acid	-	-	-	+	-	+	-	+	+	+
Bromo-succinic acid	+	-	-	-	-	-	-	+	+	+
p-hydroxy-phenyl acetic acid	-	-	-	-	+	+	-	-	-	-
Glucuronamide, α-*keto*-glutaric acid	w	-	-	-	-	+	-	-	+	+
Grow in presence of										
Acetic acid	+	-	+	+	+	+	+	+	-	+

Sodium lactate	+	+	-	+	+	+	-	+	+	+
Grow in presence of										
Fusidic acid	w	+	-	+	+	-	-	+	+	+
Lithium chloride	w	+	-	+	-	-	-	+	+	+
Potassium tellurite	+	+	+	+	+	-	-	+	+	+
Sodium bromate	w	+	+	+	-	-	-	-	+	+
Nitrogen sources										
Guanidine hydrochloride	+	+	-	+	-	-	-	+	+	-
D-serine	w	+	+	+	-	-	-	-	+	+
Antibiotic resistance to[#]										
Lincomycin	R	R	S	R	S	S	R	R	R	S
Nalidixic acid	w	R	R	R	R	S	S	R	R	R
Minocycline and vancomycin	R	R	S	R	R	S	S	R	R	R
Major fatty acids (>15%)	*iso*-$C_{16:0}$, $C_{17:1}$ ω 8c, $C_{15:0}$	*iso*-$C_{16:0}$, $C_{17:1}$ ω8c	*iso*-$C_{16:0}$, $C_{17:1}$ ω8c	*iso*-$C_{16:0}$, $C_{17:1}$ ω 8c	*iso*-$C_{16:0}$, $C_{17:1}$ ω8c	$C_{18:1}$ ω9c, $C_{16:0}$	*iso*-$C_{16:0}$, $C_{16:0}$, $C_{17:1}\omega$8c	$C_{17:1}\omega$8c , *iso*-$C_{16:0}$, $C_{16:0}$	*iso*-$C_{16:0}$, $C_{17:1}\omega$8c, $C_{17:0}$, $C_{15:0}$	*iso*-$C_{16:0}$, $C_{17:1}$ ω8c, $C_{15:0}$
Predominant menaquinones (>20%)	MK-9(H_4); MK-9(H_6)	MK-9(H_8), MK-9(H_4) [14]	MK-9(H_4), MK-9(H_6) [21]	MK-9(H_8) [15]	MK-9(H_6), MK-9(H_8) [14]	MK9(H_6) , MK9(H_4) [16]	MK-9(H_4), MK-9(H_6) [14]	MK-9(H_4) [19]	MK-9(H_6), MK-9(H_4) [20]	MK-9(H_6) [22]
Polar lipids	PI, DPG, GPL_{1-2}, PG, APL, UL	PI, DPG, GPL_{1-3}, PG, UL [14]	PI, DPG, PG, PL[21]	PI, DPG, GPL_{1-2}, PG, PL_{1-3}, UL[15]	PI, DPG, GPL_{1-3}, PG, UL[14]	PI, PG, DPG, GPL_{1-2}, UL [16]	PI, DPG, GPL_{1-3}, PG, UL[14]	PI, DPG, GPL_{1-3}, PG, UL[19]	PI, DPG, GPL_{1-2}, PG, UL[20]	PI, DPG, GPL_{1-2}, $GL_{1-6,}$ PG, PL, UL[22]

Host plant origin	*Casuarina equisetifolia*	*Alnus viridis* ssp.*crispa*	*Morella californica*	*Alnus incana ssp. rugosa*	*Casuarina cunninghamiana*	*Coriaria japonica*	*Elaeagnus angustifolia*	*Discaria trinervis*	*Elaeagnus umbellata*	*Coriaria nepalensis*
Host plant range	*Rhamnales*	*Alnus,Comptonia, Myrica*	-	*Alnus*	*Casuarinaceae* (excluding *Gymostom*), *Myricaceae*	*Coriariaceae, Datiscaceae*	*Elaeagnacea, Colletieae, Morella*	*Colletieae, Elaeagnacea eMorella*	*Elaeagnaceae, Morella*	-
Genomic G+C content (%)	70.9	72.8	72.0	72.4	70.1	71.0	71.7	72.3	72.3	71.8

+, positive reaction; -, w, weak reaction; negative reaction; R, resistant; S, sensitive; DPG: diphosphatidylglycerol; UL: unidentified lipids; PG: phosphatidylglycerol; GPL: Unknown glycophospholipid; PI: phosphatidylinositol; PL: phospholipids

Table 2. 16S rRNA gene sequence identities and dDDH values between type strain G2^T and the type strains of the nearest phylogenetic *Frankia* species. dDDH values are in % (upper right) and 16S rRNA gene sequence similarities are in % (lower left)

	G2^T	*F. inefficax* DSM 45817^T	*F. elaeagni* DSM 46783^T	*F. discariae* DSM 46785^T	*F. asymbiotica* DSM 100626^T	*F. saprophytica* DSM 105290^T
G2^T	-	22.1 [19.8 -24.6%]	25.9 [23.6 - 28.4%]	24.9 [22.6 - 27.4%]	22.5 [20.3 - 25%]	22.8 [20.5 - 25.2%]
F. inefficax DSM 45817^T	97.6	-	22.2 [20 - 24.7%]	22.6 [20.3 - 25%]	25.8 [23.5 - 28.3%]	25.7 [23.4 - 28.2%]
F. elaeagni DSM 46783^T	99.5	97.8	-	25.6 [23.3 - 28.1%]	22.5 [20.3 - 25%]	23.0 [20.7 - 25.4%]
F. discariae DSM 46785^T	99.4	97.8	98.9	-	23.1 [20.8 - 25.5%]	23.3 [21 - 25.7%]

F. asymbiotica DSM 100626^{T}	97.8	98.1	98.0	97.8	-	34.6 [32.2 - 37.1%]
F. saprophytica DSM 105290^{T}	98.1	98.5	98.2	98.0	99.4	-

[...] confidence interval

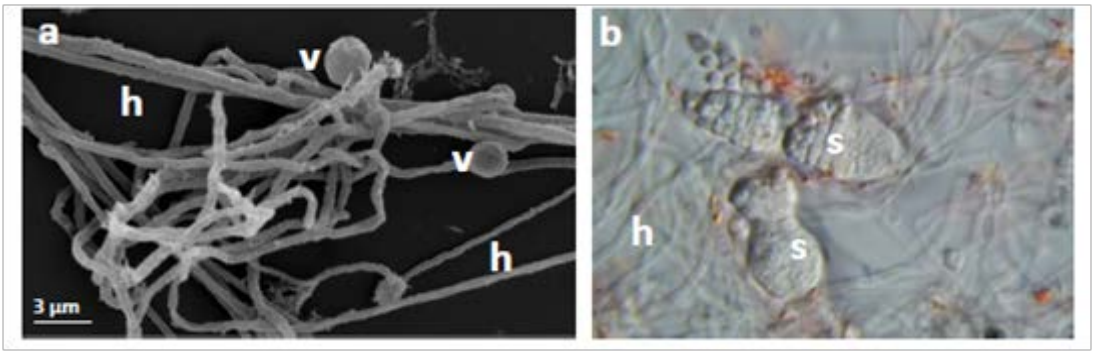

Figure 1

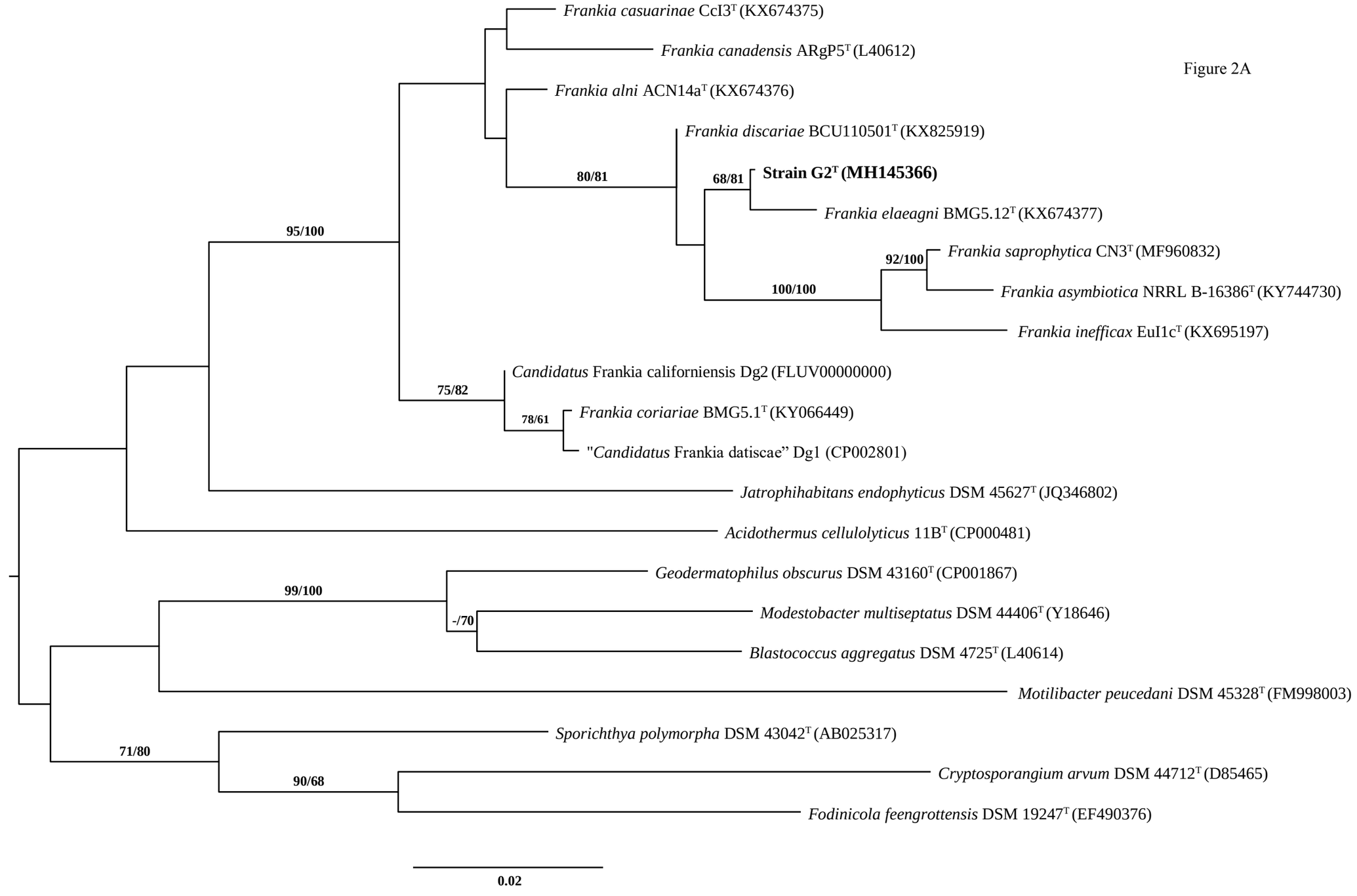
Frankia casuarinae CcI3T (KX674375)
Frankia canadensis ARgP5T (L40612)
Frankia alni ACN14aT (KX674376)
Frankia discariae BCU110501T (KX825919)
80/81
68/81
Strain G2T (MH145366)
Frankia elaeagni BMG5.12T (KX674377)
95/100
92/100
Frankia saprophytica CN3T (MF960832)
100/100
Frankia asymbiotica NRRL B-16386T (KY744730)
Frankia inefficax EuI1cT (KX695197)
Candidatus Frankia californiensis Dg2 (FLUV00000000)
75/82
78/61
Frankia coriariae BMG5.1T (KY066449)
"Candidatus Frankia datiscae" Dg1 (CP002801)
Jatrophihabitans endophyticus DSM 45627T (JQ346802)
Acidothermus cellulolyticus 11BT (CP000481)
Geodermatophilus obscurus DSM 43160T (CP001867)
99/100
Modestobacter multiseptatus DSM 44406T (Y18646)
-/70
Blastococcus aggregatus DSM 4725T (L40614)
Motilibacter peucedani DSM 45328T (FM998003)
Sporichthya polymorpha DSM 43042T (AB025317)
71/80
Cryptosporangium arvum DSM 44712T (D85465)
90/68
Fodinicola feengrottensis DSM 19247T (EF490376)
0.02

Figure 2A

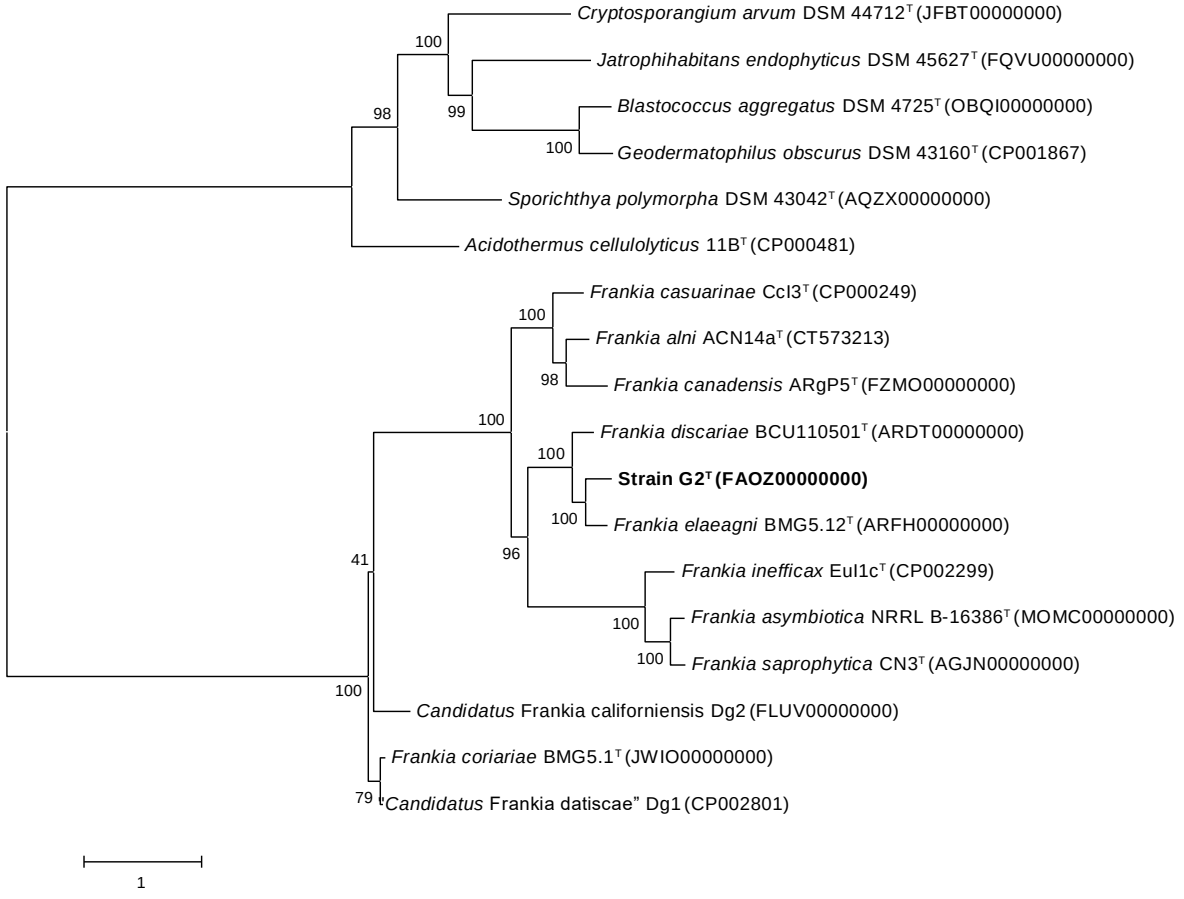
Cryptosporangium arvum DSM 44712^{T} (JFBT00000000)
Jatrophihabitans endophyticus DSM 45627^{T} (FQVU00000000)
Blastococcus aggregatus DSM 4725^{T} (OBQI00000000)
Geodermatophilus obscurus DSM 43160^{T} (CP001867)
Sporichthya polymorpha DSM 43042^{T} (AQZX00000000)
Acidothermus cellulolyticus 11B^{T} (CP000481)
Frankia casuarinae CcI3^{T} (CP000249)
Frankia alni ACN14a^{T} (CT573213)
Frankia canadensis ARgP5^{T} (FZMO00000000)
Frankia discariae BCU110501^{T} (ARDT00000000)
Strain G2^{T} (FAOZ00000000)
Frankia elaeagni BMG5.12^{T} (ARFH00000000)
Frankia inefficax EuI1c^{T} (CP002299)
Frankia asymbiotica NRRL B-16386^{T} (MOMC00000000)
Frankia saprophytica CN3^{T} (AGJN00000000)
Candidatus Frankia californiensis Dg2 (FLUV00000000)
Frankia coriariae BMG5.1^{T} (JWIO00000000)
Candidatus Frankia datiscae" Dg1 (CP002801)
100
99
98
100
100
98
100
100
100
96
41
100
100
100
79
1

Figure 2B